doi:10.3233/978-1-61499-898-3-320

Ballistocardiogram-Based Heart Rate Variation Monitoring Using Unsupervised Learning

Gang SHEN^{a,1}, Mingqi Yang^a and Biyong ZHANG^b

^aHuazhong University of Science and Technology, China ^bHangzhou Bobo Technologies, China and Eindhoven University of Technology, the Netherlands

> Abstract. Continuous heart rate variation monitoring is an important measure to diagnose and prevent the fatal cardiovascular diseases. Due to the recent development of the wearable sensory devices, ballistocardiographic (BCG) records can be collected over a long period of time without the restrictions to the patient's normal activities. However, the physical activities of the patient may severely interfere with the BCG sampling and thus leading to the degrading of the signal quality. In this paper, we introduce a novel approach to recovering the beat-to-beat heart interval estimates by applying an unsupervised machine learning framework consisting of the bag of words model and the n-gram model. First, we adopt L1 and L2 norm minimization based linear filters to reveal the global signal patterns. Then we learn the cluster centers of the signal segments containing local peaks with a Gaussian mixture approach. N-gram model is applied to detect the repetitive pattern of the cluster centers within the signals, the learned heartbeat model acts as a further filter by the convolution with the time-varying signals. Finally, autocorrelation based interval prediction is used with the help of a Gaussian prior. The proposed approach is tested using the data collected from a number of subjects with different ages and genders. We compare the resulting estimates with the blood pressure oximeter references, and the experiments show that the proposed algorithm is able to provide reliable and accurate estimates with the standard error for the interval estimate of 12.2ms.

> Keywords. ballistocardiographs, L1/L2 norm minimization, unsupervised learning

Introduction

It is well known that the cardiovascular deceases have been in the short list of the worldwide top killers for human lives over the last few decades. Long term monitoring of an individual's heart conditions has proven to be of great value for the early diagnosis of cardiovascular lesion, timely prevention of severe damage, as well as developing healthy life style. However, the traditional heart monitoring devices extensively deployed in medical practice, such as electrocardiograms (ECG), require the strict examination setup and the standard operations, in order to achieve the precision of the measurement. Attributed to this demanding scenario, patients have to visit professional medical practitioners to have their cardiovascular health evaluated.

¹ Corresponding Author, Mail: gang_shen@hust.edu.cn.

On the other hand, ballistocardiograms (BCG) render the graphical records of a person's blood ejection into the vessels powered by each heartbeat, reflecting the mechanical activities of one's cardiovascular system. In the past, it has been shown by a number of researchers that BCG can be used to observe how one's heart functions. Since BCG can be collected in a non-invasive way from the surface of a person's body, it is easily installed into other common articles such as tables, couches, beds, and weighing scales, making it a reasonable alternative to many other monitoring instruments for the purpose of everyday use. In particular, the recent advancement of electronics and wireless communications has made it possible to integrate the BCG data acquisition equipment into a wearable mobile device, allowing the continuous monitoring of the vascular system over a long period of time without causing the discomfort to the users.

The review of the recent BCG research literature can be found in [1] and [2]. The field of ballistocardiography is experiencing a resurgence due to the development of novel technologies and signal processing methods for measurement and analysis. Customized signal processing algorithms have led to reduced measurement noise, clinically relevant feature extraction, and signal modeling. The development in BCG has attracted researchers from various fields. Now there are four major types of BCG devices, depending on the way the data get collected: standing position, sitting position, prone position and wearable [3]. A variety of signal processing techniques have been applied to BCG tasks, e.g., local maximum based heart beat locating and feature extraction [4] and inverse spectrum-based heart rate calculation [5]. On top of the fast progress of the BCG data acquisition device, people have been working on understanding the physiological nature of the BCG signals. The current research is mainly focused on the correlation between the BCG signal features and cardiological diagnosis indicators, e.g. by studying a group of healthy individuals, it has been found that the time interval between the J peaks of BCG and R peaks of ECG is related to the PEPI. Also, it is shown that the consistency of BCG signals is related to the heart function from the observations on heart failure patients.

The motivation of the investigation in this paper is the detection of heart rate variability (HRV). Heart rate variability indicates the small difference between consecutive heart beats, produced by the regulation to the sinoatrial node by the autonomic nervous system. This is represented by the difference in a few dozens of milliseconds of the cardiac interphase. Though it is normal physiological behavior existing in common in cardio motility, reflecting the degree to which the sinus rhythm fluctuates under the regulation of the autonomic nervous system, HRV can be used to evaluate the cardiovascular fitness, diagnose the arrhythmia, assess one's mood, detect fatigue, and applied to the early warning of hypertension and other cardiovascular deceases.

There are many problems with BCG HRV prediction. Among the most remarkable challenges that BCG applications are facing is the signal quality variation. Due to the lack of standard it is highly dependent on the way a person uses the BCG data acquisition device, the signal varies greatly if the contact of the device to the body surface gets loose or tight. Because the BCG senses the force instead of directly measuring bioelectric signals, it may be compromised by the physical movement of one's body and other interferences. Distinguished with the existing research, first, we prose to learn a L2 norm based linearized filter that associates the signal fragments far from the current sample to reveal the global pattern of the BCG observation. Second, we propose a clustering algorithm that groups the similar signal segments into a given

number of classes, before these classes are further processed to detect the repetitive patterns; finally, we adopt the Viterbi algorithm to align the signal with the selected pattern in order to mark each heart beat to avoid the impact of the corrupt signal sections.

The remainder of this paper is organized as follows: in Section 2, we review the related work in the fields of BCG applications, including the BCG signal processing techniques and heart rate detection methods proposed by other researchers; in Section 3, we present the main algorithms of this paper; in Section 4, we test the proposed approach with the BCG data collected from a group of health young adults, and compare our approach with other algorithms to demonstrate the effectiveness of our proposal.

1. Related work

During the resurgence of the unobtrusive monitoring of cardiovascular activities in health-care practice in the last decade, the research community has paid attention to various BCG signal sources, especially the signals collected from the sleeping position. Also the state-of-the-art signal processing methods like wavelet filters have been applied in order to deal with BCG signals in absence of the quality guarantee [6]. Machine learning algorithms have been modified to fit the needs arising from BCG signal processing, in particular, clustering [7] and Bayesian approach [8] are used to find the BCG signal pattern related to heart beat. After the pattern is learned, autocorrelation and dynamic programming can be used for the beat-to-beat interval detection [9] [10].

In particular, multiple methods have been recently proposed for the detection of individual beat-to-beat intervals. As an important quantitative measure of cardiovascular regulation by the autonomic nervous system, beat-to-beat heart rate variability bears significantly more information than the heart rate averaged over time. Heart beat model is a valuable indicator to detect the individual heartbeat and has drawn the interest of many scientists. As reported in literature, when tested on different sizes of subject groups, though the detection coverage of beat-to-beat intervals may vary in a noticeable degree, the mean error of the beat-to-beat intervals falls within a couple of dozens of milliseconds in most cases, compared with the benchmark ECG results [11].

Different heart beat modes have been analyzed. Once the model is obtained, BCG signals can be used to match the model in order to find the maximal local correlation. Since the medical professionals have the in-depth observation on the physiological behavior of cardiac system, the template can be manually defined by experts. In [12], the authors considered a fiducial peak point of BCG as an I-J-K complex which corresponds with ventricle contraction and Electrocardiogram (ECG) QRS complex. They proposed a beat detection algorithm consisting of two stages, template definition stage and beat detection stage with defined template in previous stage. In the first stage, the BCG template is constructed by the expert with an empirical analysis of BCG signal and measurement device. In the second stage, the correlation function calculates an accuracy of template with BCG signal using a local moving window.

Alternatively, another way to define the heart beat model is to apply different machine learning techniques. In [13] and [14], targeting the home-use e-health system for an unobtrusive sleep measurement, the authors presented a model for extracting the

waveform of the heartbeat from the BCG by using unsupervised learning techniques such as hierarchical clustering. Then, beat-to-beat intervals are detected by finding positions where the heartbeat shape best fits the signal.

In [15], the authors presented an algorithm for the detection of individual heart beats in ballistocardiograms (BCGs). In a training step, unsupervised learning techniques are used to identify the shape of a single heart beat in the BCG. The learned parameters are combined with so-called "heart valve components" to detect the occurrence of individual heart beats in the signal. A refinement step improves the accuracy of the estimated beat-to-beat interval lengths. In [16], the authors presented a method for the precise detection of heartbeats from a ballistocardiography (BCG) signal. First, feature vectors are extracted from the signal at possible heartbeat positions. Clustering is then applied to the vectors to find a cluster with the highest density. The densest cluster corresponds to the most accurately repeating shape in the signal and thus the positions of the feature vectors of the densest cluster should match real heartbeat positions in the signal. In [17] the authors designed algorithms to compute heart rate and assess heart rate variability from the perspective of the structure of human heartbeats, by modeling the stochastic structure of heartbeat intervals as an inverse Gaussian process, attempting to estimate heart rate variability in real time.

Noise reduction is a necessary preprocessing step before handling BCG signals. In [18], the authors focused on the tact that the BCG devices still suffer from the interferences induced due to subject movement or motion during signal acquisition or even due to floor vibrations. Artifact removal is necessary for efficient analysis and health monitoring. The authors proposed a method of signal processing, while the raw signal is pre-processed and parsed to independent component analysis which provides the decomposed components and later k-means is applied to detect the components which are responsible for artifact and removed.

2. Proposed approach

In this section, we discuss the proposed approach for BCG signal based HRV with the intended application into healthcare industry to monitor the everyday cardiovascular status of the patients without any negative impact to their normal behavior. The BCG data is collected from the PVdF sensors placed under the seating and the signal quality is not always satisfactory for the purpose of further processing. In the proposed framework, we divide the algorithms into three phases: preprocessing, pattern detection and heartbeat interval analysis.

2.1. Data collection

We use PVdF to collect the BCG data for sitting position. Piezoelectric thin films are a sort of newly developed high polymer energy exchanging material with the unique dielectric effect, piezoelectric effect and thermoelectric effect. Compared with the conventional piezoelectric materials (such as ceramics) it bears the characteristics of wide frequency response, large dynamic range, high sensitivity, sensitive force to electricity conversion as well as the mechanical properties of high strength, and acoustic impedance matching. PVdF also has the advantages of being light-weight, soft and non-crispy, impact resistant, less susceptible to water and chemical pollution, and can be easily shaped in different pieces. However, since the sensors are mounted

underneath the cushion, and the patients may not sit still for very long period of time in a day, the BCG signals suffer from external interferences such as the patients physical move and the vibration of the environment. Therefore, the challenge is to reduce the noise's impact to the handing of BCG signals in order to find the robust prediction of HRV.

2.2. Preprocessing of BCG signals

We apply the second order Butterworth to exclude most portion of the signals unrelated to the cardiovascular behavior. To be specific, we first take the Butterworth high pass filter with the cut-off frequency of 1Hz, in order to get rid of the low frequency effect of breath. Then we use the low pass Butterworth filter with cut off frequency of 10Hz, to exclude the high frequency noise (in Figure 1, we can see that after filtering, the remaining BCG signals preserve the components closely associated with the cardiovascular features).



Figure 1. The signals before (up) and after the Butterworth filtering.

Nevertheless, the remaining BCG signal now still look very different both in magnitude and in waveform. In order to make the signal more consistent over time, global information covering a longer time span helps to reshape the signal. From this perspective, we design the linear filter to reveal the global pattern in the signal. Since the signal is considered a time series, the signal value at a point is closely related to the signal values around that point. Assuming the normal heartbeat follows the same signal pattern, we can approximate this pattern by finding the signal's linear correlation at a point with a fragment of signals centered at that point, and this linear relationship is a candidate for the pattern representation. However, since the strong local similarity between the signal at a pint and the signals adjacent to that point, we need to exclude those signal points in the neighborhood of the point under consideration, to make sense of the global property. Both L1 and L2 norm constraints on the linear coefficients are able to provide insights on the filter design. If the signals are repetitive in a fragment, minimizing L2 norm may be more useful. On the other hand, L1 norm minimization finds the sparse representation of the linear relationship, and this is valuable in the sense that the heartbeat signal is sparse in essence. We present a two-round filtering procedure, first a L2 norm-based filter is used to find the signal estimate, then a second L1 norm filter is applied to generate the sparse approximation of the signal. First, we

take a batch normalization to the signals as a necessary step to reduce the inconstancy in signal strengths, since the sensors can be very sensitive to the varying pressure from a sitting person. In the filter's objective function, both L2 and L1 norms of the filter coefficients are to be minimized while allowing the signal to be approximated with a given error bound.

$$\min ||a||_i, i = \{1, 2\}$$

subject to $\sum_{\tau} ||\mathbf{y}_{\tau} - a^T x_{\tau}||_2 \le \epsilon, \tau \in T$

where T is a time period, x_{τ} is the signal value at time instant $\tau \in T$, and the vector $\mathbf{y}_{\tau} = \{x_t | t \in (\tau + T_1, \tau + T_2) | | t \in (\tau - T_1, \tau + T_2) \}$, and finally a is the design parameter to be optimized. The selection of correlation band width is not as straightforward as it first looks, we suggest that the length of the signal fragment $2T_2$ better cover a whole heartbeat interval while the dead zone width better be similar to half of the BCG signal peak width.

In addition, we adopt sequential k-means clustering (k=2) to exclude the abnormal signal segments reflecting the body movement (high in variance) and loose contact (low in signal variation) respectively.

2.3. Pattern learning

The pattern is learned using a bag of words model that is widely found is textual documents mining. We treat the time varying signal as a set of contiguous segments that contain single peaks and thus partition the signal by troughs, i.e., the local minimums of the signal. The portion of signal that sits between two consecutive troughs are then recorded as length varying vectors. These vectors are subsequently cut into fixed length vectors by keeping only the portion surrounding the peak, say, with the width of n points. After the set of fixed length vectors are available, Gaussian mixture method is applied to cluster the vectors into a given number of clusters. The underlying assumption is that the vectors are randomly generated from a few fixed components of a normal heartbeat, with similar vectors being drawn from the same component.

After the clusters are found, we assign the cluster names to each of the signal segments mentioned before. These segments form a text consisting of the symbols representing the cluster names. Then a n-gram model is used to find the cooccurrences of the symbols. In the case of BCG signals, we only use 2-gram to calculate the transition frequencies of the symbols.

Starting from any symbol, we follow the most probable transitions to the next symbol, until we find the repetition. This gives the repetitive template of symbols that most likely represents the heartbeat pattern.

The final step within the pattern learning stage is to retrieve the heartbeat signal pattern out of the signals, based on the pattern of symbols as just learned. This is achieved by averaging all portions of the signals that match the symbolic template, as shown in Figure 2. Suppose the peaks of a signal segment following the heartbeat pattern are denoted $(pk_1, pk_2, ..., pk_n)$ the average heartbeat interval is estimated as

$$h = \frac{len(pk_1pk_2) + len(pk_2pk_3) + \dots + len(pk_{n-1}pk_n)}{m}$$

As the result, we will use this signal pattern to find the beat-to-beat interval estimates.



Figure 2. Signal segements (left) and the learned pattern.

After the heartbeat pattern is learned, we use that pattern to calculate the convolution with the BCG signals (see Figure 3), the resulting time series make the heart beats more uniform and ready for the prediction of heart rate or beat-to-beat intervals.

$$\Phi_p[m] = \frac{1}{m|p||s|} \sum_{v=0}^m p[v]s[v-m]$$

where p is the pattern vector and s is the signal segment vector.



Figure 3. Convolution computed using the learned pattern before (up) and after pattern.

2.4. Beat-to-beat interval estimate

Subsequently, we apply autocorrelation to estimate the beat-to-beat intervals [10]. The length of the segment that maximizes the auto-correlation should be considered a proper estimate for the beat-to-beat interval. Due to the randomness in the nature of the cardiovascular physiology, the signal can deviate from the pattern in both the magnitude and width, thus making the optimum not an obvious winner in practice, because there could be a lot of lengths with the similar correlation results due to the

existence of many artifacts. The adaptive correlation is basically the inner product of two vectors. We normalized the result with the vectors lengths.

$$\Phi_{adapt}[m] = \frac{1}{m} \sum_{v=0}^{m} w[v]w[v-m]$$

2.5. Gaussian prior in estimating the interval

Since the auto-correlation based beat-to-beat interval estimate is compromised by various artifacts, in particular, at some instants, the artifacts may be more correlated than the true heartbeat, in this paper, we propose a Gaussian prior by assuming that the heartbeats at consecutive time instants follow a Gaussian distribution. This prior can effectively avoid the consequences caused by artifacts (as shown in Figure 4).



Figure 4. Auto-correlation spectrum before (left) and after Gaussian prior Then the local maximum corresponds to the beat-to-beat interval estimate.

3. Experiments and analysis

We collected the BCG data using PVdF for 10 test subjects in 2017 using the sampling rate of 125Hz. Among the subjects, there are 5 males and 5 females, with age between 25 to 35 years old. No Cardiovascular diseases have been reported for them. In the mean time, the finger tip blood pressure oximeters were used for photoplethysmography (PPG) reference. These records will be used to find the ground truth. Each of the subjects was tested in seated position for a few minutes. Signals lasting 1 minutes were stored for further analysis.

As discussed before, the signals were first preprocessed with the Butterworth filter and the L1 and L2 norm based linear filters. Then the signal segments containing 15 points before and after a time instant were used for clustering. After the heartbeat model is learned, we may piece together the correlation calculations needed to estimate the intervals. The instant interval estimates apply to any point on the time axis, thus the variability of the heartbeat interval can be detected instantly.



Figure 5. Auto-correlation spectrum over time (left) and estimated intervals.

We compare the estimates with the ground truths obtained from the reference PPG signals. Since BCG and PPG differ in phase, we manually adjust the starting points of both signals to match their phase. Then the estimate errors can be found and used to evaluate the performance of the proposed approach. We can see that the estimate generated by the proposal (blue) follows the PPG ground truth (red) closely after the introduction of the Gaussian prior.

We tested different setups of the Gaussian prior parameters, it turned out that if the variance stays in a range centered at $\sigma = 30$, the best estimation is achieved (see the left of Figure 6, the estimation is in value while the benchmark is in red). However, if the variance goes up to 100, there will be a few sudden spikes in the estimation as shown in Figure 6 (right). The robustness is evidenced by the experiments, as we know that all beat-to-beat intervals were covered by the proposed approach.



Figure 6. Comparison of different Gaussian prior variances.

The learned heartbeat pattern plays an important role in reshaping the waveform of the processed BCG signals. Without the filtering effect from the pattern, the standard estimation error of a single interval averaged for all test subjects is 25.1ms, and after the pattern filtering, the standard error is lowered to 12.2ms (varying from 11.15ms to 12.27ms), which outperforms the results reported in other work.

4. Concluding remarks

In this paper, we presented our work on improving the beat-to-beat interval estimation from BCG signals with low quality. Since there could be many artifacts introduced during the data acquisition, filtering is a necessary step to remove the noise and preserve the cardiovascular features. We proposed basically two sorts of filters: first we apply the constrained L1 and L2 norm minimization based linear filters to make the global properties more prominent over the signals; then we take the unsupervised learning methods to find the heartbeat model, to be specific, we cut the signal into peak segments and cluster these segments to find the cluster centers with Gaussian mixture model, and the transition frequencies with the help of n-gram model, and consequently the learned model can be used as a pattern filter. The beat-to-beat interval is finally estimated using the auto-correlation computation, with a Gaussian prior being added on top of the auto-correlation results to exclude the impact of artifacts, and thus reducing the interval error to 12ms in the experiments, compared with the reference signals. The test results show that the proposed approach is able to produce reasonably good estimate for daily HRV monitoring.

References

- [1] O.T. Inan, P.F. Migeotte, K.S. Park, M. Etemadi, K. Tavakolian, R. Casanella, J. Zanetti, J. Tank, I. Funtova, G.K. Prisk and M. Di Rienzo, Ballistocardiography and seismocardiography: a review of recent advances, *IEEE Journal of Biomedical & Health Informatics*, 2015, Vol. 19(4), pp. 1414-1427.
- [2] L. Giovangrandi, O.T. Inan, R.M. Wiard, M. Etemadi and G.T. Kovacs, Ballistocardiography A Method Worth Revisiting, *Engineering in Medicine and Biology Society, Embc, 2011 International Conference of the IEEE*, IEEE, 2011, DOI: 10.1109/IEMBS.2011.6091062.
- [3] D. Friedrich, X.L. Aubert, H. Fuhr and A. Brauers, Heart rate estimation on a beat-to-beat basis via ballistocardiography - a hybrid approach, *Engineering in Medicine and Biology Society*, IEEE, 2010, pp. 4048-4051.
- [4] B.H. Choi, G.S. Chung, J.S. Lee, et al. Slow-wave sleep estimation on a load-cell-installed bed: a nonconstrained method, *Physiological Measurement*, Vol. 30(11), 2009, pp. 1163-1170.
- [5] J.M. Kortelainen and J. Virkkala, FFT averaging of multichannel BCG signals from bed mattress sensor to improve estimation of heart beat interval, *Engineering in Medicine and Biology Society*, 2007. *Embs 2007, International Conference of the IEEE*, IEEE, 2007, pp. 6685-6688.
- [6] Q. Delière, J. Tank, I. Funtova, E. Luchitskaya, D. Gall, P. Van de Borne and P.-F. Migeotte, Ballistocardiogram amplitude modulation induced by respiration: a wavelet approach, 2015 Computing in Cardiology Conference (CinC), Nice, 2015, pp. 313-316.
- [7] L. Rosales, M. Skubic, D. Heise, M.J. Devaney and M. Schaumburg, Heartbeat detection from a hydraulic bed sensor using a clustering approach, *Engineering in Medicine & Biology Society. Conf Proc IEEE Eng Med Biol Soc*, 2012, pp. 2383-2387.
- [8] J. Paalasmaa, A respiratory latent variable model for mechanically measured heartbeats, *Physiological Measurement*, Vol. 31(10), 2010, pp. 1331-1344.
- [9] C. Brüser, S. Winter and S. Leonhardt, Robust inter-beat interval estimation in cardiac vibration signals, *Physiological Measurement*, Vol. 34(2), 2013, pp. 123-138.
- [10] C. Brüser, S. Winter and S. Leonhardt, How speech processing can help with beat-to-beat heart rate estimation in ballistocardiograms, *IEEE International Symposium on Medical Measurements and Applications Proceedings*, IEEE, 2013, pp. 12-16.
- [11] S.K. Jose, C.M. Shambharkar and J. Chunkath, HRV analysis using ballistocardiogram with LabVIEW, International Conference on Computing and Communications Technologies. IEEE, 2015, pp. 128-132.
- [12] J.H. Shin, B.H. Choi, Y.G. Lim, et al. Automatic ballistocardiogram (BCG) beat detection using a template matching approach, *Engineering in Medicine and Biology Society*, 2008. Embs 2008. International Conference of the IEEE, IEEE, 2008, pp. 1144-1146.
- [13] J. Paalasmaa, H. Toivonen and M. Partinen, Adaptive Heartbeat Modelling for Beat-to-beat Heart Rate Measurement in Ballistocardiograms, *IEEE Journal of Biomedical & Health Informatics*, Vol. 19(6), 2015, pp. 1945-1952.
- [14] C. Bruser, K. Stadlthanner, S.D. de Waele and S. Leonhardt, Adaptive Beat-to-Beat Heart Rate Estimation in Ballistocardiograms, *IEEE Transactions on Information Technology in Biomedicine*, Vol. 15(5), 2011, pp. 778-786.
- [15] C. Brüser, K. Stadlthanner, A. Brauers and S. Leonhardt, Applying machine learning to detect individual heart beats in ballistocardiograms, *International Conference of the IEEE Engineering in Medicine & Biology. Conf Proc IEEE Eng Med Biol Soc*, 2010, pp. 1926-1929.
- [16] J. Paalasmaa and M. Ranta, Detecting heartbeats in the ballistocardiogram with clustering, Proceedings of the ICML/UAI/COLT 2008 Workshop on Machine Learning for Health-Care Applications, Helsinki, Finland, 2008.
- [17] R. Barbieri, E.C. Matten, A.A. Alabi and E.N. Brown, A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability, *American Journal of Physiology Heart* & Circulatory Physiology, Vol. 288(1), 2005, pp. 424-35.
- [18] B.M. Manjula and C. Sharma, BCG Artifact Removal Using Improved Independent Component Analysis Approach, *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 5(1), 2017, pp. 130.